PLoS one

QM/MM MD and Free Energy Simulations of G9a-Like Protein (GLP) and Its Mutants: Understanding the Factors that Determine the Product Specificity

Yuzhuo Chu[1,2], Jianzhuang Yao[1], Hong Guo[1,3]*

1 Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee, United States of America, 2 National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, Tennessee, United States of America, 3 University of Tennessee/Oak Ridge National Laboratory Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee, United States of America

Abstract

Certain lysine residues on histone tails could be methylated by protein lysine methyltransferases (PKMTs) using S-adenosyl-L-methionine (AdoMet) as the methyl donor. Since the methylation states of the target lysines play a fundamental role in the regulation of chromatin structure and gene expression, it is important to study the property of PKMTs that allows a specific number of methyl groups (one, two or three) to be added (termed as product specificity). It has been shown that the product specificity of PKMTs may be controlled in part by the existence of specific residues at the active site. One of the best examples is a Phe/Tyr switch found in many PKMTs. Here quantum mechanical/molecular mechanical (QM/MM) molecular dynamics (MD) and free energy simulations are performed on wild type G9a-like protein (GLP) and its F1209Y and Y1124F mutants for understanding the energetic origin of the product specificity and the reasons for the change of product specificity as a result of single-residue mutations at the Phe/Tyr switch as well as other positions. The free energy barriers of the methyl transfer processes calculated from our simulations are consistent with experimental data, supporting the suggestion that the relative free energy barriers may determine, at least in part, the product specificity of PKMTs. The changes of the free energy barriers as a result of the mutations are also discussed based on the structural information obtained from the simulations. The results suggest that the space and active-site interactions around the ε-amino group of the target lysine available for methyl addition appear to among the key structural factors in controlling the product specificity and activity of PKMTs.

Introduction

The tails of histone proteins are subject to a variety of post-translational modifications, and these modifications are believed to be a part of the histone code for chromatin regulation [1]. One of such modifications is the methylation of a range of lysine (K) residues on histones, including K4, K9, K27, K36, K79 on histone H3 and K20 on histone H4 [2]. Catalyzed by protein lysine methyltransferases (PKMTs), histone lysine methylations have been found to affect a variety of important biological processes, including heterochromatin formation, X-chromosome inactivation, transcriptional silencing and activation [3,4]. PKMTs may be classified based on their ability to transfer one, two or three methyl groups from S-adenosyl-L-methionine (AdoMet, the methyl donor) to the ε-amino group of target lysine, a special property of the enzymes that is termed as product specificity [1,4]. Since different methylation states may lead to different downstream events [5–7], it is of fundamental importance to understand the

determinant of the product specificity, including the energetic and structural factors that control the product specificity.

A number of experimental [8–17] and computational studies [18–26] have been performed to understand the origin of the product specificity for PKMTs. Structural and mutational studies have identified a Phe/Tyr switch located at the active site of many SET domain PKMTs [9,10,12,14]. It has been shown previously that the product specificity of the enzyme may depend on whether this position is occupied by a Phe or Tyr residue. Indeed, it has been observed that this position tends to be occupied by a tyrosine residue for mono-methyltransferases (e.g. Y305 in SET7/9 [17] and Y334 in SET8 [10,13]) and by a phenylalanine for di- or tri-methyltransferases (e.g. F281 in DIM-5 [9]). Moreover, the substitution at the Phe/Tyr switch position could lead to the change of the product specificity that is consistent with the observations on the wild-type enzymes (see above). For example, the F281→Y mutation for DIM-5 altered the enzyme from a tri-methyltransferase to a mono-/di-methyltransferase [9], while the Y305→F mutation for SET7/9 [17] or Y334→F mutation for

SET8 [10,13] was found to switch the enzyme from a mono-methyltransferase to a di-methyltransferase. For H3K9 G9a methyltransferase, the F→Y mutant at the Phe/Tyr switch site (F1152 in G9a) was found to alter the enzyme from a di-methyltransferase to a mono-methyltransferase [14]; the equivalent residue in its closely relative G9a-like-protein (GLP) is F1209. Previous studies on SET7/9, SET8 and DIM-5 [13,17–19] have suggested that this Phe/Tyr switch may affect product specificity through altering the affinity of an active site water molecule, as the dissociation of this water molecule is likely to make it easier for further methyl addition [13,17–19]. It would be of interest to examine whether this is the case for GLP or G9a as well.

The alternations of the product specificity of PKMTs by single mutations have also been observed for residues located at some other positions of the active site [9,15,17], although detailed studies for the origin of the change of the product specificity due to these mutations are still lacking. For GLP there is a Tyr residue (Y1124) hydrogen-bonding to the methyl accepting nitrogen in the X-ray structures of GLP complexed with the substrate, H3K9me and H3K9me2 peptides [15]; the equivalent residue in G9a is Y1067. It has been found that the Y1067→F mutant for G9a is able to increase the enzyme's ability to add a third methyl group to the target lysine (even though it is not at the Phe/Tyr switch position) and change the enzyme to a tri-methyltransferase [15]. Further studies are still necessary to determine the origin of the product specificity change for this case.

Hybrid quantum mechanical/molecular mechanical (QM/MM) free energy simulations have been performed previously on various SET domain PKMTs, including DIM-5, SET7/9, SET8 and their mutants involving the residue at the Phe/Tyr switch site [18–20]. One of the main purposes was to study the energetic origin of the product specificity and identify the structural factors which might control the methylation states of target lysine residues [18–20]. For instance, it has been proposed that the existence of a relatively high free energy barrier for one or more methyl transfer processes could stop the continuation of the methyl addition and may therefore determine how the epigenetic marks of lysine methylation are written [18–20]. However, questions remain whether the previous conclusions could be extended to mutations involving the residues other than those at the Phe/Tyr

Figure 1. The definition of the structural parameters for monitoring the relative orientation of AdoMet and H3K9me1 [H3K9 and H3K9(me)$_2$] in the reactant complex. The efficiency of the methyl transfer may be related to the distributions of $r(C_M...N_\zeta)$ and θ in the reactant complexes. θ is defined as the angle between the two vectors r_1 and r_2. Here r_1 is the direction of the lone pair of electrons on N_ζ and r_2 is the vector pointing from C_M to S_δ. The reaction coordinate for calculating the free energy profiles for the methyl transfers is $R = r(C_M...S_\delta) - r(C_M...N_\zeta)$.
doi:10.1371/journal.pone.0037674.g001

switch site. In this paper, the results of the QM/MM molecular dynamics (MD) and free energy simulations are reported for different methyl transfer processes of wild type (WT) GLP as well as its F1209Y and Y1124F mutants. The free energy barriers of the methyl transfer processes obtained from simulations are well correlated with experimental observations of product specificities [14,15], supporting previous conclusions that the relative free energy barrier for the methyl transfer process may determine, at least in part, the number of methyl groups added by the enzyme [18–20]. The dynamic information of reactant complexes gathered from the QM/MM MD simulations were also examined, and the results suggest that the ability of the reactant complex to form the reactive configuration may be used as an important indicator for the product specificity of PKMTs.

Methods

The QM/MM MD and free energy (potential of mean force) simulations were applied to determine the free energy profiles of different methyl transfer processes catalyzed by WT GLP, F1209Y and Y1124F mutants. MD simulations were also performed on the reactant complexes of the methyl transfers for each system to study their dynamic properties related to the methyl transfers. AdoMet and the target lysine side chains were treated by QM and the rest of the system were treated by MM. The link-atom approach [27] was implemented with CHARMM [28] to separate the QM and MM regions. A modified TIP3 water model [29,30] was used for the solvent. The stochastic boundary molecular dynamics method [31] was used for the QM/MM MD and free energy simulations. The system was separated into a reaction zone and a reservoir zone, and the reaction zone was further divided into a reaction region and a buffer region. The reaction region was a sphere with radius r of 20 Å, and the buffer region had r equal to 20 Å$\leq r \leq$22 Å. The reference center for partitioning the system was chosen to be the C_δ of the target K-9. The resulting systems contained around 5500 atoms, including about 800–900 water molecules. There is a crystal water molecule in the active site of the X-ray structure, and no additional water at the active site was introduced through the solvation process.

The SCC-DFTB method [32] implemented in CHARMM [28] was used for the QM atoms and the all-hydrogen CHARMM potential function (PARAM27) [33] was used for the MM atoms. In our earlier studies, the results of the SCC-DFTB and B3LYP/6-31G** methods for the description of the methyl transfer in a small model system were compared using energy minimization-based approach [18–20]. This comparison allowed us to understand the performance of the semi-empirical method in the description of the bond breaking and making and derive an empirical correction for the free energy curves obtained from the potential of mean force simulations. The results indicated that the energy curves from the corrected SCC-DFTB and B3LYP/6-31G** were very close, supporting the use of this approach to make the first correction to the bond breaking and making events involving the simple and similar SN2 methyl transfer processes. The similar approach of this "first-order" correction to the free energy curves was also adopted in this work. For a more detailed discussion, see Ref.20. It should be pointed out that cares must be exercised when applying this procedure to study other methyl transfer processes involving different methyl donors and acceptors (such as the methyl transfer in protein arginine methyltransferases). In such cases, a model system containing the corresponding methyl donor and acceptor should be used instead for determining the deficiency of the SCC-DFTB method in the description of methyl transfer. The results of SCC-DFTB calculations on the methyl transfer for
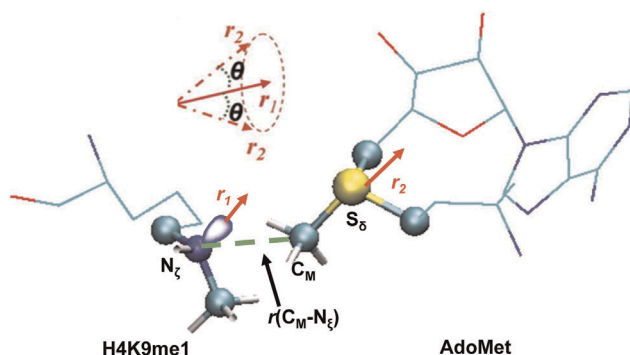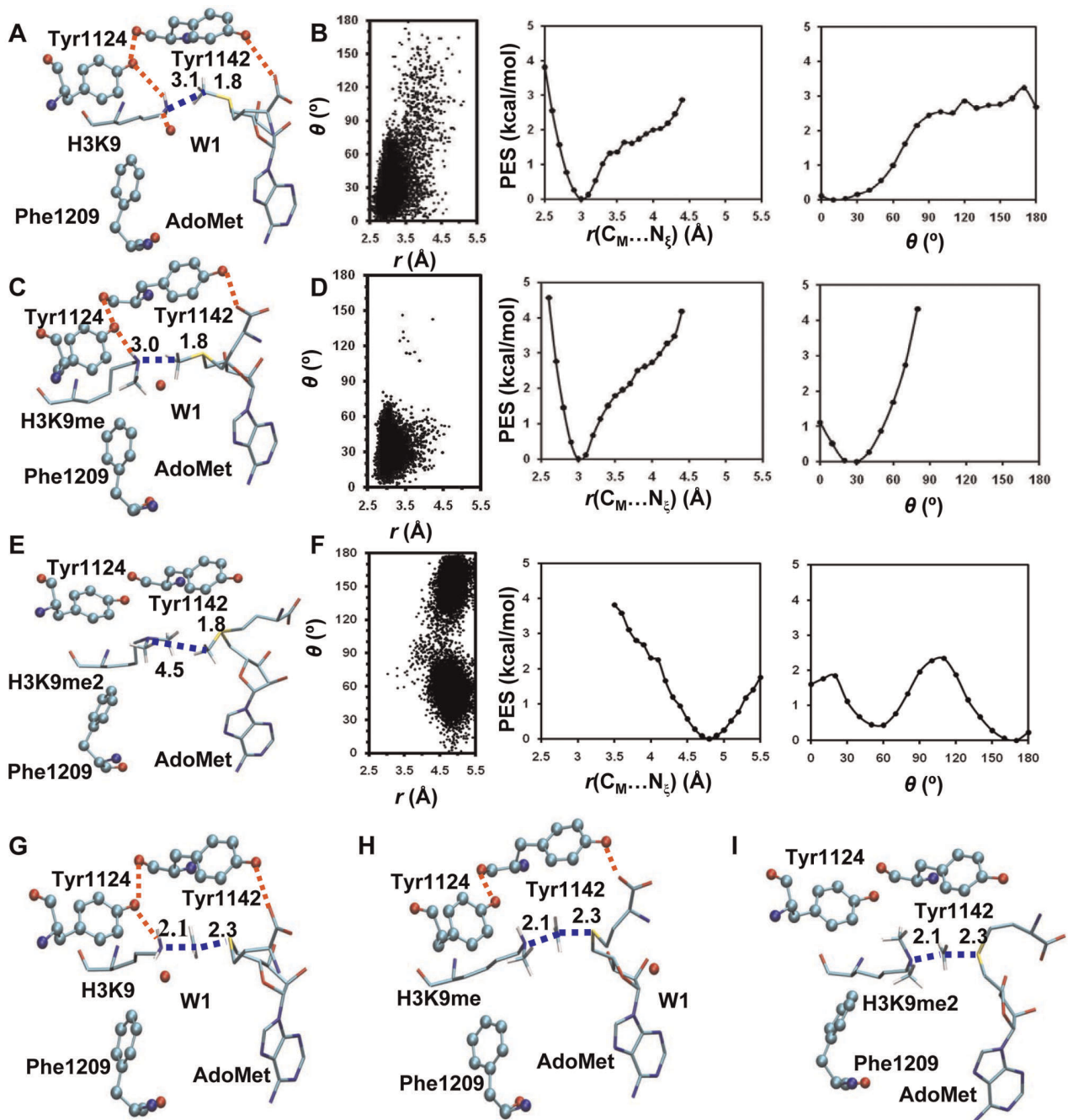
Figure 2. MD results for the WT enzyme (GLP). (A) A representative active-site structure along with the average values of some structural parameters of the reactant complex for the first methyl transfer. GLP is shown in balls and sticks, and AdoMet and the H3K9 sidechain are in sticks. Hydrogen atoms are not shown for clarity except for those on N_ζ and transferable methyl group. Hydrogen bonds are indicated by red dotted lines, and the distances related to the reactant coordinate are also shown. (B) Left: the two-dimensional plot of $r(C_M \ldots N_\zeta)$ and θ distributions based on the 1.5-*ns* simulations of the reactant complex for the first methyl transfer; Middle: the free-energy change as a function of $r(C_M \ldots N_\zeta)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (C) The active-site structure along with the average values of some structural parameters of the reactant complex for the second methyl transfer. (D) Left: the two-dimensional plot of $r(C_M \ldots N_\zeta)$ and θ distributions of the reactant complex for the second methyl transfer; Middle: the free-energy change as a function of $r(C_M \ldots N_\zeta)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (E) The structure of the reactant complex for the third methyl transfer. (F) Left: the two-dimensional plot of $r(C_M \ldots N_\zeta)$ and θ distributions of the reactant complex for the third methyl transfer; Middle: the free-energy change as a function of $r(C_M \ldots N_\zeta)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (G) A representative active-site structure along with the average values of some structural parameters near the transition state for the first methyl transfer obtained from the free energy (potential of mean force) simulations. (H) The structure along with the average values of some structural parameters near the transition state for the second methyl transfer. (I) The structure near the transition state for the third methyl transfer. All images were made by VMD [38]. The distances shown on the structures are the calculated average distances from the trajectories over the 50-ps production runs in the corresponding window.
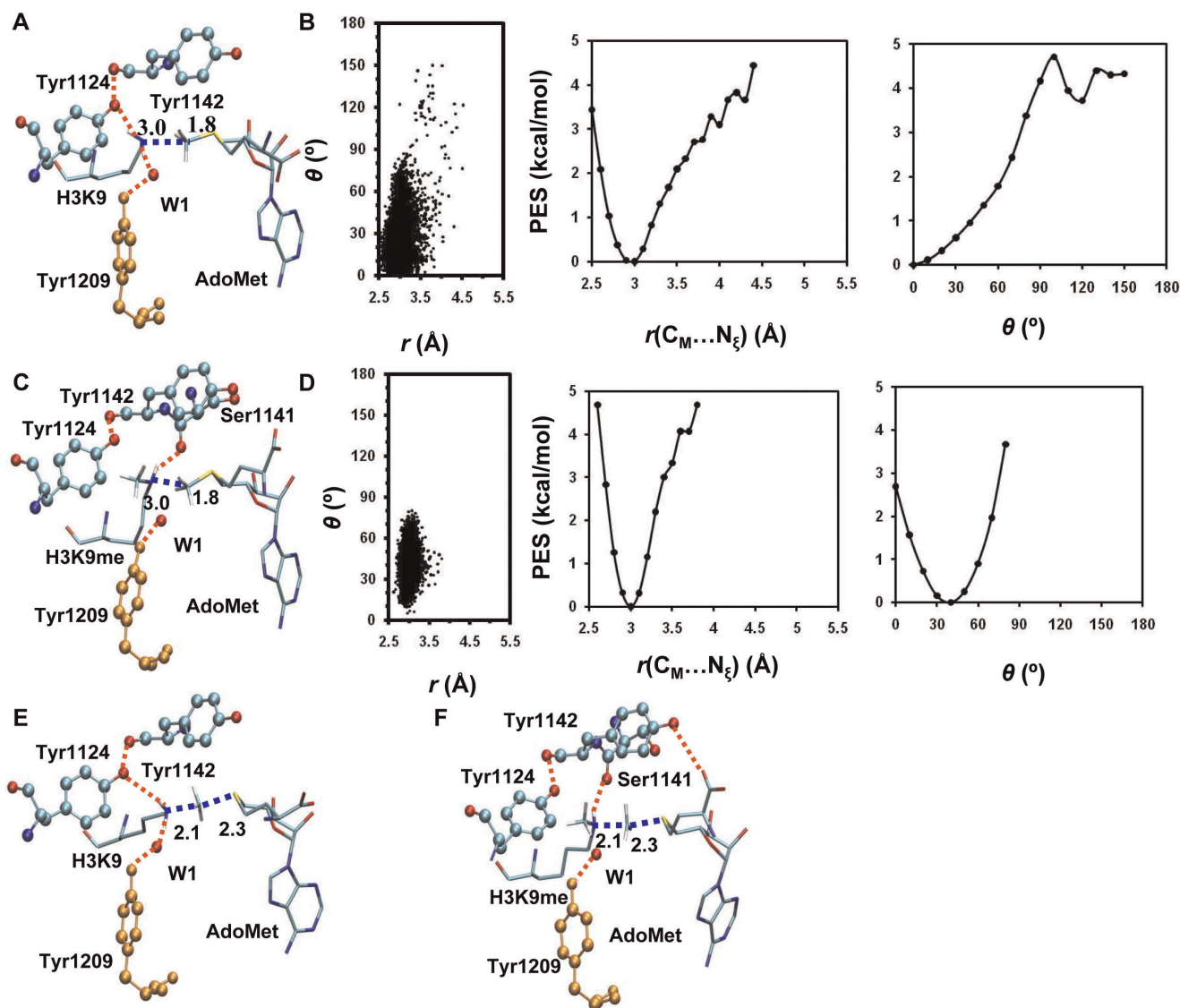doi:10.1371/journal.pone.0037674.g002

Figure 3. MD results for the F1209Y mutant. (A) A representative active-site structure along with the average values of some structural parameters of the reactant complex for the first methyl transfer. (B) Left: the two-dimensional plot of $r(C_M...N_\zeta)$ and θ distributions of the reactant complex for the first methyl transfer; Middle: the free-energy change as a function of $r(C_M...N_\zeta)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (C) The structure of the reactant complex for the second methyl transfer. (D) Left: the two-dimensional plot of $r(C_M...N_\zeta)$ and θ distributions of the reactant complex for the second methyl transfer; Middle: the free-energy change as a function of $r(C_M...N_\zeta)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (E) The structure near the transition state for the first methyl transfer. (F) The structure near the transition state for the second methyl transfer.
doi:10.1371/journal.pone.0037674.g003

such model systems can be compared with the results from high-level ab initio (e.g., B3LYP/6-31G**) calculations to obtain reasonable corrections for such methyl transfer process.

The initial coordinates for the reactant complexes of the first, second and third methyl transfers were based on the crystallographic complexes (PDB codes: 2RFI, 3HNA) of GLP containing AdoHcy and short H3K9me1 or H3K9me2 peptide [15]. A methyl group was added to AdoHcy to form AdoMet. The initial coordinates for F1209Y and Y1124F mutants were generated by changing the corresponding amino acid manually. For the first methyl transfer step, methyl group in H3K9me1 was removed to generate H3K9. The initial structures for the entire stochastic boundary system were optimized using the steepest descent (SD) and adopted-basis Newton-Raphson (ABNR) methods. The

systems were gradually heated from 50.0 to 283.15 K in 50 ps and equilibrated at 283.15 K for 500 ps. A 1-fs time step was used for integration of the equation of motion, and the coordinates were saved every 50 fs for analyses. After 500-ps of equilibration were performed, 1-ns QM/MM MD simulations were carried out for each of the reactant complexes of the first, second and third methyl transfers. As discussed in the previous studies [18–20], the S_N2 methyl transfer from AdoMet to K9, K9me1 or K9me2 is presumably more efficient if the S-CH$_3$ group of AdoMet is well aligned with the lone pair of the electrons of N_ζ in the reactant complex; i.e., with a small θ angle and relatively short C_M-N_ζ distance. Therefore, we determined the distributions of $r(C_M$-$N_\zeta)$ and θ from the MD trajectories; θ was defined as the angle
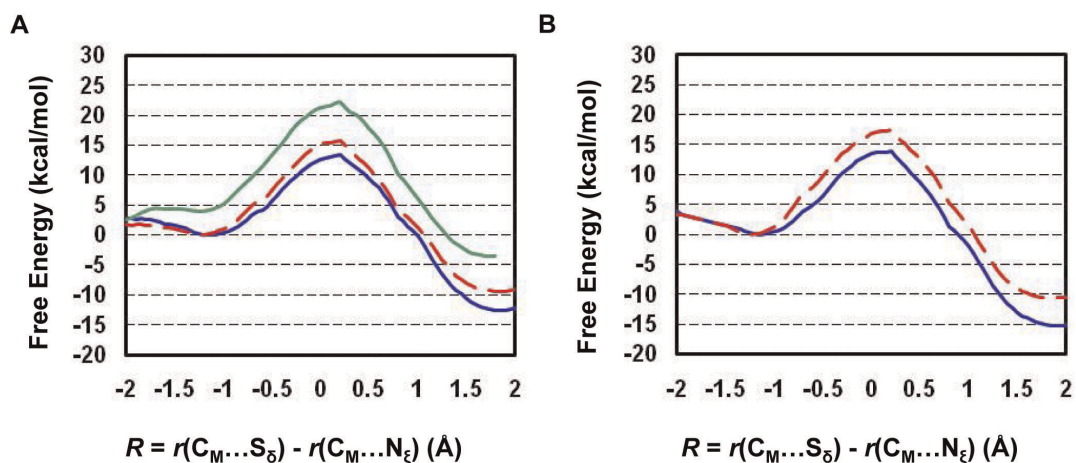
A

B

Figure 4. Free energy profiles of methyl transfer processes in WT GLP and the F1209Y mutant. (A) The free energy (potential of mean force) changes for the first, second and third methyl transfers from AdoMet to H3-K9, H3-K9me1 and H3-K9me2, respectively, as a function of the reaction coordinate $[R = r(C_M...S_\delta) - r(C_M...N_\xi)]$ in the wild-type GLP. The first methyl transfer: blue and solid line with a free energy barrier of 13.4 kcal/mol; the second methyl transfer: red and dashed line with a free energy barrier of 15.8 kcal/mol (or about 2.4 kcal/mol higher than that of the first methyl transfer); the third methyl transfer: green and solid line with a free energy barrier of 22.1 kcal/mol (or about 8.7 kcal/mol higher than that of the first methyl transfer). (B) The free energy changes for the first and second methyl transfers as a function of the reaction coordinate in the F1209Y mutant. The first methyl transfer: blue and solid line with a free energy barrier of 13.8 kcal/mol; the second methyl transfer: red and dashed line with a free energy barrier of 17.4 kcal/mol (or about 3.6 kcal/mol higher than that of the first methyl transfer).
doi:10.1371/journal.pone.0037674.g004

between the direction of the C_M-S_δ bond (r_2) and the direction of the electron lone pair on N_ξ (r_1) (see **Figure 1**).

The umbrella sampling method [34] implemented in the CHARMM program along with the Weighted Histogram Analysis Method (WHAM) [35] was applied to determine the change of the free energy (potential of mean force) as a function of the reaction coordinate for the methyl transfer from AdoMet to H3-K9, H3-K9me1 or H3-K9me2 in the WT and mutated enzyme. The reaction coordinate was defined as a linear combination of $r(C_M$-$N_\xi)$ and $r(C_M$-$S_\delta)$ $[R = r(C_M$-$S_\delta) - r(C_M$-$N_\xi)]$ (see **Figure 1**). For each methyl transfer process, twenty windows were used, and for each window 50-ps production runs were performed after 20-ps equilibration. The force constants of the harmonic biasing potentials used in the PMF simulations were from 50 to 400 kcal mol^{-1} Å^{-2}. More sophisticated simulation methods [36,37] may be used in the future to generate the free energy surfaces of methyl transfers that also depend on angular coordinates.

Results

The active-site structure along with the average values of some structural parameters of the reactant complex for the first methyl transfer in WT GLP is shown in **Figure 2A**. As is clear from this figure, the lone pair of electrons on N_ξ of the target lysine is well aligned with the methyl group of AdoMet in the structure, a condition that was found to make the methyl transfer relatively easier to occur. The distribution plot from the MD simulations for the reactant complex (**Figure 2B**) also shows that there are large populations of the structures with relatively short $r(C_M$-$N_\xi)$ distance and small values of the θ angle; the average $r(C_M$-$N_\xi)$ distance is around 3.0 Å and the average θ angle is about 30°. **Figure 2A and 2G** (the structure near the transition state obtained from the free energy simulations along with the average values of some structural parameters) show that Tyr1124 and an active-site water molecule (W1) form stable hydrogen bonds with the ε-amino group of target lysine, and these hydrogen bonds may help to orientate the electron lone pair and achieve a good

alignment with the methyl donor group of AdoMet. The active-site structure along with the average values of some structural parameters of the reactant complex for the second methyl transfer in WT GLP is given in **Figure 2C**. Similar to the case for the first methyl transfer, a good alignment between the lone pair of electrons and the methyl group from AdoMet can also be achieved in this reactant complex (also see **Figure 2D**). Examination of the structure in **Figure 2C** shows that while Tyr1124 still forms a hydrogen bond with the ε-amino group of target lysine as in the case for the first methyl transfer, the hydrogen bond with W1 disappears as a result of losing the hydrogen-bond donating capacity for the methylated lysine. It is of interest to note from the average structure near the transition state (**Figure 2H**) that W1 has dissociated from the active site during the methyl transfer. The dissociation of this water presumably produces additional space in the active site that makes the second methyl transfer easier. The active-site structure along with the average values of some structural parameters of the reactant complex for the third methyl transfer is shown in **Figure 2E**. As is demonstrated in this figure, the lone pair of electrons cannot be well aligned with the methyl group; the average $r(C_M$-$N_\xi)$ distance is about 4.5 Å and the values of θ angle become significantly larger with a broad distribution (**Figure 2F**).

The active-site structure along with the average values of some structural parameters of the reactant complex for the first methyl transfer in the F1209Y mutant is shown in **Figure 3A**. As is the case for the wild type enzyme, the lone pair of electrons on N_ξ of target lysine is well aligned with the methyl group of AdoMet in the reactant complex for the first methyl transfer. **Figure 3C and 3D** show that for the second methyl transfer, the chances for the θ angle to reach 0–15° (i.e., near the line arrangement) become smaller in the mutant than in wild-type. Moreover, the free energy cost for generating the linear configuration is as much as 2–3 kcal/mol (only about 1 kcal/mol in wild-type, see **Figure 2D**). This energy cost may be related to the relatively higher free energy barrier for the second methyl transfer in F1209Y (see below). The structure near the transition state (**Figure 3F**) shows that unlike in
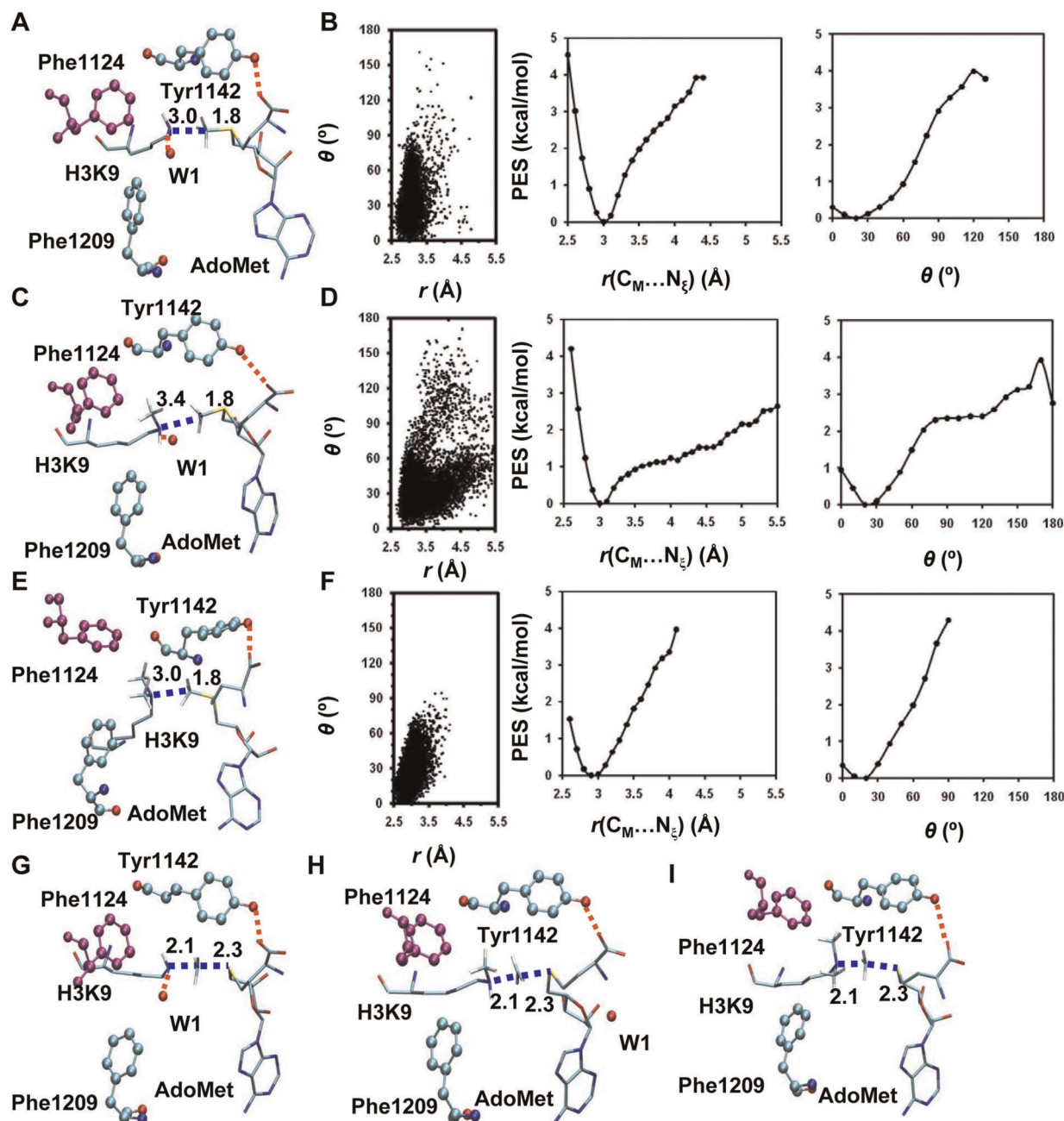
Figure 5. MD results for the Y1124F mutant. (A) A representative active-site structure along with the average values of some structural parameters of the reactant complex for the first methyl transfer. (B) Left: the two-dimensional plot of $r(C_M…N_\zeta)$ and θ distributions of the reactant complex for the first methyl transfer; Middle: the free-energy change as a function of $r(C_M…N_\xi)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (C) The structure of the reactant complex for the second methyl transfer. (D) Left: the two-dimensional plot of $r(C_M…N_\zeta)$ and θ distributions of the reactant complex for the second methyl transfer; Middle: the free-energy change as a function of $r(C_M…N_\xi)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (E) The structure of the reactant complex for the third methyl transfer. (F) Left: the two-dimensional plot of $r(C_M…N_\zeta)$ and θ distributions of the reactant complex for the third methyl transfer; Middle: the free-energy change as a function of $r(C_M…N_\xi)$ obtained from the distributions; Right: the free-energy change as a function of θ obtained from the distributions. (G) A representative active-site structure along with the average values of some structural parameters near the transition state for the first methyl transfer. (H) The structure near the transition state for the second methyl transfer. (I) The structure near the transition state for the third methyl transfer.
doi:10.1371/journal.pone.0037674.g005

the case of wild-type, W1 remains in the active site during the second methyl transfer. The failure for the dissociation of W1 from the active site is presumably due to the presence of the hydrogen bond between W1 and Y1209 (that stabilizes the water molecule),

and the existence of W1 is likely to make it more difficult for further methyl addition.

The free energy profiles for methyl transfer processes in WT GLP and the F1209Y mutant are plotted in **Figure 4A and 4B**, respectively, as a function of the reaction coordinate
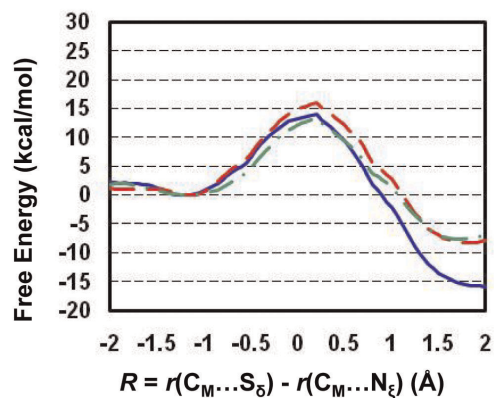
Figure 6. Free energy profiles of methyl transfer processes in the Y1124F mutant. The first methyl transfer: blue and solid line with a free energy barrier of 13.9 kcal/mol; the second methyl transfer: red and dashed line with a free energy barrier of 15.9 kcal/mol; the third methyl transfer: green and dashed line with a free energy barrier of 13.3 kcal/mol.
doi:10.1371/journal.pone.0037674.g006

$[R = r(C_M...S_\delta) - r(C_M...N_\zeta)]$. Figure 4A shows that the increase of the free energy barrier from the first or second methyl transfer to that of the third methyl transfer in WT GLP is considerably large (6–9 kcal/mol), consistent with the fact that WT GLP is a di-methylatransferase [14]. For the F1209Y mutant, there is a significant increase of the free energy barrier from the first to second methyl transfer (3.6 kcal/mol) and the chance for observing dimethyl lysine from F1209Y is expected to be much smaller compared to that for wild-type. The results are consistent with the experimental observation that F1209Y mutation altered the enzyme from a di-methyltransferase to a mono-methyltransferase [14]. These data agree with the previous conclusions that the existence of a relatively high free energy barrier could stop the further addition of methyl group and may therefore control the product specificity of PKMTs [18–20], at least in many cases.

The active site structures along with the average values of some structural parameters of the reactant complexes for the first, second and third methyl transfer of the Y1124F mutant are shown in **Figure 5A, 5C, 5E**, respectively. All the structures have the lone pair of electrons on N_ζ of target lysine well aligned with the methyl donor group of AdoMet. The distribution plots for these structures also show large population of structures with relatively short $r(C_M-N_\zeta)$ distances and small values of the θ angle; i.e. the average $r(C_M-N_\zeta)$ distances are around 3.0 Å and the average angles are in the range of 0–30° (**Figure 5B, 5D, 5F**). The distributions of the $r(C_M-N_\zeta)$ distance and θ angle for the second methyl transfer is broader compared with that observed in wild type enzyme. This could be caused by the loss of the stabilization from the hydrogen bond involving Tyr1124. Nevertheless, the free energy costs for generating the linear configuration is rather small (**Figure 5D**). The substitution of Tyr1124 by Phe may be able to lead to additional space around the ε-amino group of the target lysine and make it easier to add the third methyl group to target lysine. Indeed, the free energy profiles of all three methyl transfer steps in the Y1124F mutant in **Figure 6** show that all the three methyl transfers have similar and relatively low free energy barriers, consistent with the experimental observation that the Y1124F mutant is a tri-methyltransferase [15].

Discussion

The results of the simulations show that the structures of the reactant complexes for the first and second methyl transfers in WT (**Figures 2A, 2C**), the first methyl transfer in F1209Y (**Figure 3A**), and the first, second and third methyl transfers in Y1124F (**Figures 5A, 5C, 5E**) are rather similar to the corresponding structures near transition state (TS). The distribution plots (**Figure 2B, 2D, 3B, 5B, 5D, 5F**) of these structures show large population of structures with relatively short $r(C_M-N_\zeta)$ distances and small values of the θ angle, which are close to the distribution plots (**Figure S1**) of the TS structures. This is in contrast to the case of the third methyl transfer in wild-type and the second methyl transfer in F1209Y mutant where the reactant structures (**Figure 2E, 3C**) are significantly distorted from the corresponding TS structures (**Figure 2I, 3F**). The distribution plots (**Figure 2F, 3D**) of these reaction complexes show large population of structures with either relatively long $r(C_M-N_\zeta)$ distances or large values of the θ angle, which are considerably different from the corresponding distribution plot (**Figure S1**) of the TS structures. Thus, the results are consistent with the previous suggestion [18–20] that one of the reasons for the existence of the relatively low free energy barrier is likely owed to the fact that a part of TS stabilization is already reflected on the reactant state through the generation of the TS-like conformation. Thus, a good alignment of the electron lone pair on N_ζ of target lysine with the methyl group of AdoMet is likely to lead to a relatively efficient methyl transfer, while a poor alignment could make the methyl transfer more difficult.

Two free energy triplets, $(0, \Delta_{2\text{-}1W}, \Delta_{3\text{-}1W})$ and $(\Delta_{M\text{-}W}, \Delta_{2\text{-}1M}, \Delta_{3\text{-}1M})$ have been proposed [18–20] to describe the energetics of the methyl transfers for PKMTs and their mutants, respectively, and to determine their product specificity. Here the free energy barrier of the first methyl transfer step in WT enzyme is taken as the zero. For WT enzyme, the second $(\Delta_{2\text{-}1W})$ and third $(\Delta_{3\text{-}1W})$ parameters are the differences in the free energy barriers between the second and first and between the third and first methyl transfers, respectively. These two parameters therefore provide information about the relative efficiencies of the second and third methyl transfers compared to the first methyl transfer. For instance, if $\Delta_{2\text{-}1W}$ is significant (e.g., in the case of SET8 [19]), the enzyme may stop the methylation process leading to a mono-methylase. For the mutated enzyme, the first parameter $(\Delta_{M\text{-}W})$ is the difference in the free energy barriers for the first methyl transfer in the WT and mutant. The second $(\Delta_{2\text{-}1M})$ and third $(\Delta_{3\text{-}1M})$ parameters are the differences in the free-energy barriers between the second and first and between the third and first methyl transfer, respectively (e.g., defined similar to the cases for WT). Based on such definitions, the energy triplets for WT, F1209Y and Y1124F mutants of GLP are (0, 2.4, 8.7), (0.4, 4.0, x) and (0.5, 2.5, −0.1), respectively (**Figure 4 and 6**). The significant increase of the free energy barrier from the first or second methyl transfer to that of the third methyl transfer in WT GLP is consistent with the observation that the enzyme is a di-methylase based on the simulations. For the F1209Y mutant, the free energy barrier of the second methyl transfer is around 3.6 kcal/mol higher than the first methyl transfer, and the mutant is more likely to be mono-methylase. It is of interest to note that the free energy barriers of all three methyl transfers in the Y1124F mutant are similar, suggesting that the replacement of Y1124 by F changes the enzyme from a di-methylase into a tri-methylase. All these results are consistent with previous experimental observations [14,15].

The structural information generated from the simulations shows that the dissociation of W1 might occur during the second methyl transfer in both WT GLP and the Y1124F mutant (**Figure 2H, 5H**). For the second methyl transfer in the F1209Y mutant, however, W1 cannot dissociate from the active site, presumably, due to its formation of the hydrogen bond with Tyr1209. Consequently, the free energy barrier increases significantly for the addition of the second methyl group, leading to a mono-methylase. These results are consistent with the earlier suggestion that the dissociation of the active site water molecule is important for further addition of methyl group to target lysine [13,17–19]. For the Y1124F mutant, additional space in the active site may be generated by the removal of the hydroxyl group on Y1124, making it easier for the addition of the third methyl group.

Conclusion

In this study, the QM/MM MD and free energy simulations have been performed on different methyl transfer steps from AdoMet to target lysine/methyl-lysine in WT GLP as well as its F1209Y and Y1124F mutants. The reactant complexes for different methyl transfers generated from QM/MM MD simulations support the suggestion that the ability of the reactant complex to form the TS-like reactive configuration (in which the electron lone pair on N_ξ of target lysine is well aligned with the methyl group of AdoMet) is important for the methyl addition. The free energy barriers of methyl transfers were also calculated. The results show that the relative free energy barriers for the three methyl transfers may determine the product specificity for WT GLP and its mutants. The role of the active site water molecule in controlling product specificity was also discussed.

Supporting Information

Figure S1   The two-dimensional plots of $r(C_M...N_\xi)$ and θ distributions of the transition state complexes. A: plots for WT GLP, left: first methyl transfer; middle: second methyl transfer; right: third methyl transfer. B: plots for F1209Y mutant, left: first methyl transfer; right: second methyl transfer. C: plots for Y1124F mutant, left: first methyl transfer; middle: second methyl transfer; right: third methyl transfer. All distribution plots are based on the 50-ps production runs of the transition state complexes.
(TIFF)

Author Contributions

Conceived and designed the experiments: YZC HG. Performed the experiments: YZC. Analyzed the data: YZC. Contributed reagents/materials/analysis tools: JZY. Wrote the paper: YZC HG.

References

1. Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403: 41–45.
2. Marmorstein R (2003) Structure of SET domain proteins: a new twist on histone methylation. Trends Biochem Sci 28: 59–62.
3. Jenuwein T (2006) The epigenetic magic of histone lysine methylation. FEBS J 273: 3121–3135.
4. Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol 6: 838–849.
5. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struc Mol Biol 14: 1025–1040.
6. Lall S (2007) Primers on chromatin. Nat Struc Mol Biol 14: 1110–1115.
7. Turner BM (2005) Reading signals on the nucleosome with a new nomenclature for modified histones. Nat Struc Mol Biol 12: 110–112.
8. Zhang X, Tamaru H, Khan SI, Horton JR, Keefe LJ, et al. (2002) Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell 111: 117–127.
9. Zhang X, Yang Z, Khan SI, Horton JR, Tamaru H, et al. (2003) Structural basis for the product specificity of histone lysine methyltransferases. Mol Cell 12: 177–185.
10. Couture JF, Collazo E, Brunzelle JS, Trievel RC (2005) Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes Dev 19: 1455–1465.
11. Xiao B, Jing C, Kelly G, Walker PA, Muskett FW, et al. (2005) Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev 19: 1444–1454.
12. Xiao B, Jing C, Wilson JR, Walker PA, Vasisht N, et al. (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. Nature 421: 652–656.
13. Couture JF, Dirk LMA, Brunzelle JS, Houtz RL, Trievel RC (2008) Structural origins for the product specificity of SET domain protein methyltransferases. Proc Natl Acad Sci U S A 105: 20659–20664.
14. Collins RE, Tachibana M, Tamaru H, Smith KM, Jia D, et al. (2005) In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. J Biol Chem 280: 5563–5570.
15. Wu H, Min JR, Lunin VV, Antoshenko T, Dombrovski L, et al. (2010) Structural Biology of Human H3K9 Methyltransferases. Plos One 5: e8570.
16. Xu ST, Wu J, Sun BF, Zhong C, Ding JP (2011) Structural and biochemical studies of human lysine methyltransferase Smyd3 reveal the important functional roles of its post-SET and TPR domains and the regulation of its activity by DNA binding. Nucleic Acids Res 39: 4438–4449.
17. Del Rizzo PA, Couture JF, Dirk LMA, Strunk BS, Roiko MS, et al. (2010) SET7/9 Catalytic Mutants Reveal the Role of Active Site Water Molecules in Lysine Multiple Methylation. J Bio Chem 285: 31849–31858.
18. Xu Q, Chu YZ, Guo HB, Smith JC, Guo H (2009) Energy Triplets for Writing Epigenetic Marks: Insights from QM/MM Free-Energy Simulations of Protein Lysine Methyltransferases. Chem Eur J 15: 12596–12599.
19. Chu YZ, Xu Q, Guo H (2008) Understanding Energetic Origins of Product Specificity of SET8 from QM/MM Free Energy Simulations: What Causes the

Stop of Methyl Addition during Histone Lysine Methylation? J Chem Theory Comput 6: 1380–1389.
20. Guo HB, Guo H (2007) Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. Proc Natl Acad Sci U S A 104: 8797–8802.
21. Zhang XD, Bruice TC (2008) Product specificity and mechanism of protein lysine methyltransferases: Insights from the histone lysine methyltransferase SET8. Biochemistry 47: 6671–6677.
22. Bai QF, Shen YL, Yao XJ, Wang F, Du YP, et al. (2011) Modeling a New Water Channel That Allows SET9 to Dimethylate p53. Plos One 6: e19856.
23. Hu P, Wang S, Zhang Y (2008) How do SET-domain protein lysine methyltransferases achieve the methylation state specificity? Revisited by ab initio QM/MM molecular dynamics simulations. J Am Chem Soc 130: 3806–3813.
24. Wang SL, Hu P, Zhang YK (2007) Ab initio quantum mechanical/molecular mechanical molecular dynamics simulation of enzyme catalysis: The case of histone lysine methyltransferase SET7/9. J Phys Chem B 111: 3758–3764.
25. Hu P, Zhang YK (2006) Catalytic mechanism and product specificity of the histone lysine methyltransferase SET7/9: An ab initio QM/MM-FE study with multiple initial structures. J Am Chem Soc 128: 1272–1278.
26. Zhang X, Bruice TC (2008) Enzymatic mechanism and product specificity of SET-domain protein lysine methyltransferases. Proc Natl Acad Sci U S A 105: 5728–5732.
27. Field MJ, Bash PA, Karplus M (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. J Comput Chem 11: 700–733.
28. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1983) CHARMM – a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4: 187–217.
29. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparision of simple potential functions for simulating liquid water. J Chem Phys 79: 926–935.
30. Neria E, Fischer S, Karplus M (1996) Simulation of activation free energies in molecular systems. J Chem Phys 105: 1902–1921.
31. Brooks CL, Brunger A, Karplus M (1985) Active-site dynamics in protein molecules – a stochastic boundary molecular-dynamics approach. Biopolymers 24: 843–865.
32. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. J Phys Chem B 105: 569–585.
33. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102: 3586–3616.
34. Torrie GM, Valleau JP (1974) Monte-carlo free-energy estimates using non-boltzmann sampling – application to subcritical lennard -jones fluid. Chem Phys Lett 28: 578–581.
35. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. J Comput Chem 13: 1011–1021.

36. Glowacki DR, Paci E, Shalashilin DV (2009) Boxed Molecular Dynamics: A Simple and General Technique for Accelerating Rare Event Kinetics and Mapping Free Energy in Large Molecular Systems. Journal of Physical Chemistry B 113: 16603–16611.

37. Ricchiuto P, Brukhno AV, Paci E, Auer S (2011) Communication: Conformation state diagram of polypeptides: A chain length induced alpha-beta transition. Journal of Chemical Physics 135.

38. Humphrey W DA, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14: 27–38.